ADA033298

THE CERTAINTY ILLUSION

OREGON RESEARCH INSTITUTE

Paul Slovic
Baruch Fischhoff
Sarah Lichtenstein

DDC
RECEIVED
DEC 8 1976
C

ADVANCED DECISION TECHNOLOGY PROGRAM

CYBERNETICS TECHNOLOGY OFFICE
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
Office of Naval Research • Engineering Psychology Programs

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

BEST
AVAILABLE COPY

The objective of the Advanced Decision Technology Program is to develop and transfer to users in the Department of Defense advanced management technologies for decision making. These technologies are based upon research in the areas of decision analysis, the behavioral sciences and interactive computer graphics. The program is sponsored by the Cybernetics Technology Office of the Defense Advanced Research Projects Agency and technical progress is monitored by the Office of Naval Research – Engineering Psychology Programs. Participants in the program are:

Decisions and Designs, Incorporated
The Oregon Research Institute
Perceptronics, Incorporated
Stanford University
The University of Southern California

Inquiries and comments with regard to this report should be addressed to:

Dr. Martin A. Tolcott
Director, Engineering Psychology Programs
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

or

LT COL Roy M. Gulick, USMC
Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

The views and conclusions contained in this document are those of the author(s) and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government. This document has been approved for public release with unlimited distribution.

TECHNICAL REPORT DDI-2

THE CERTAINTY ILLUSION

by

Paul Slovic, Baruch Fischhoff,
and Sarah Lichtenstein

Sponsored by

Defense Advanced Research Projects Agency
Contract N00014-76-C-0074
ARPA Order No. 3052
Under Subcontract from
Decisions and Designs, Incorporated

August, 1976

DDC
RECEIVED
DEC 8 1976
C

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

OREGON RESEARCH INSTITUTE
P.O. Box 3196
Eugene, Oregon 97403
(503) 484-2123

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

SUMMARY

Introduction

The purpose of this study is to demonstrate that the class of facts that we know with certainty is smaller than we think it is. That is, many of the opinions to which we give the privileged status of being "certain" or "almost certain" fact are incorrect. We shall refer to this phenomenon as "the certainty illusion."

Background and Approach

Subjects in four experiments answered a variety of general knowledge questions and indicated their degree of certainty about each answer. The correctness of those answers about which they were certain provided a test of the certainty illusion.

Findings and Implications

Subjects were frequently wrong on answers they judged certain to be correct. Careful tutoring of subjects in the subtleties of expressing their certainty in terms of probabilities and odds did little to reduce the illusion. Feelings of certainty were so strong that subjects were willing to bet on the correctness of their knowledge. Because of the illusion, the bets they accepted were quite disadvantageous to them. The psychological basis for unwarranted certainty is discussed in terms of the inferential processes whereby knowledge is reconstructed from fragments of perceptions and memories.

Feelings of certainty often lead to bold, decisive action. If what we know with certainty is, in fact, untrue, our resultant acts could produce catastrophic outcomes. By alerting decision makers to this phenomenon and helping them understand why it occurs, the incidence of unwarranted and potentially destructive certainties may be reduced.

TABLE OF CONTENTS

	Page
SUMMARY	ii
LIST OF TABLES	iv
ACKNOWLEDGMENT	v
INTRODUCTION	1
EXPERIMENT 1	1
METHOD	1
RESULTS	4
EXPERIMENT 2	6
METHOD	6
RESULTS	8
EXPERIMENT 3	8
METHOD	10
RESULTS	10
EXPERIMENT 4	11
METHOD	11
RESULTS	12
GAMBLING FOR MONEY	16
CONCLUSION	17
APPENDIX: DETAILED INSTRUCTIONS FOR EXPERIMENTS 3 AND 4	25
REFERENCES	31
DISTRIBUTION LIST	32
DD FORM 1473	35

LIST OF TABLES

Table		Page
1.	Lethal Events Whose Relative Frequencies were Judged in Experiment 1.	3
2.	Percentage of Correct Answers for Major Odds Categories.	5
3.	Analysis of Certainty Responses in Experiment 2.	9
4.	Odds Responses to Five High Certainty-Low Accuracy Items in Experiment 4.	23

ACKNOWLEDGMENT

Support for this research performed by Oregon Research Institute was provided by the Advanced Research Projects Agency of the Department of Defense and was monitored under Contract N00014-76-C-0074 with the Office of Naval Research under subcontract from Decisions and Designs, Inc.

THE CERTAINTY ILLUSION

INTRODUCTION

"Certainty generally is an illusion, and repose is not the destiny of man."

--Oliver Wendell Holmes, The Path of Law, 1897.

Have you ever ventured an opinion on a factual matter about which you felt, "I couldn't possibly be wrong!"? If you're like us, you have this feeling often. There are some facts that are so obvious to you that you would be insulted if their correctness were questioned--facts that you feel you know, "as well as I know my own name." After all, if we could not trust ourselves to know some things with certainty, how could we function in this world? The purpose of this study is to demonstrate that the class of facts that we know with certainty is smaller than we think. That is, many of the opinions that we give the privileged status of being "certain" or "almost certain" fact are incorrect. We shall refer to this phenomenon as "the certainty illusion."

Data are presented from four experiments in which subjects answered a variety of general knowledge questions. The correctness of those answers about which they were certain provided the test of the certainty illusion. In the first experiment, the stimuli were pairs of lethal events. For each pair, subjects were asked to estimate the more frequent event and then to indicate their degree of certainty about the correctness of their judgment.

EXPERIMENT 1

Method

Stimuli. The first experiment studied the accuracy with which people could judge the relative frequencies of the 41 lethal events shown in Table 1.

These events were chosen as a representative subset of the larger set of lethal events for which yearly statistics are available. The events in Table 1 are ordered according to frequency of death from each cause per 10^8 United States residents per year. Event frequencies were estimated from recent vital statistics reports, primarily those prepared by the National Center for Health Statistics and the Statistical Bulletin of the Metropolitan Life Insurance Company. These frequencies provided the basis for the correct answers to the questions put to our subjects.

From among these 41 causes of death, 106 question-pairs were constructed so that each cause appeared in approximately six pairs and the ratios of the statistical rates of the more frequent event to the less frequent event varied systematically from 1.25:1 (e.g., accidental falls vs. emphysema) to about 1,000,000:1 (e.g., stroke vs. botulism).

Procedure. Subjects' instructions read as follows:

> Each item consists of two possible causes of death. The question you are to answer is: which cause of death is more frequent, in general, in the United States?
>
> For each pair of possible causes of death, A and B, we want you to mark on your answer sheet which cause you think is more frequent.
>
> Next, we want you to decide how confident you are that you have, in fact, chosen the more frequent cause of death. Indicate your confidence by the odds that your answer is correct. Odds of 2:1 mean that you are twice as likely to be right as wrong. Odds of 1,000:1 mean that you are a thousand times more likely to be right than wrong. Odds of 1:1 mean that you are equally likely to be right or wrong. That is, your answer is completely a guess.

Table 1

Lethal Events Whose Relative Frequencies were Judged in Experiment 1. Death rate per 10^8 U. S. residents per year is shown to the left of each event.

0	Smallpox
0.5	Poisoning by vitamins
1	Botulism
2.4	Measles
3	Fireworks
4	Smallpox vaccination
7.2	Whooping cough
8.3	Polio
23.5	Venomous bite or sting
44	Tornado
52	Lightning
63	Non-venomous animal
100	Flood
163	Excess cold
200	Syphilis
220	Pregnancy, childbirth and abortion
330	Infectious hepatitis
440	Appendicitis
500	Electrocution
740	Motor vehicle-train collision
920	Asthma
1100	Firearm Accident
1250	Poisoning by solid or liquid
1800	Tuberculosis
3600	Fire and flames
3600	Drowning
7100	Leukemia
8500	Accidental falls
9200	Homicide
10,600	Emphysema
12,000	Suicide
15,200	Breast cancer
19,000	Diabetes
27,000	Motor vehicle (car, truck or bus) accident
37,000	Lung cancer
46,400	Cancer of the digestive system
55,000	All accident
102,000	Stroke
160,000	All cancer
360,000	Heart disease
849,000	All disease

Use any odds you wish. For example, you could write 75:1 if you think that it is 75 times more likely that you are right than that you are wrong, or 1.2:1 if you think your chances of being correct are only slightly greater than your chances of being incorrect.

Do not use odds less than 1:1. That would mean that it is less likely that you are right than that you are wrong, in which case you should indicate the other cause of death as more frequent.

In case some of the causes of death are ambiguous or not well defined by the brief phrase that describes them, we have included a glossary for several of these items. Read this glossary before starting.

Subjects. The subjects were 66 paid volunteers who answered an ad in the University of Oregon student newspaper.

Results[1]

The appropriateness of subjects' odds estimates can be assessed by looking at what is called their "degree of calibration" (Lichtenstein, Fischhoff & Phillips, in press). If subjects were "well calibrated," we should find that they were correct on about 50% of the answers for which they gave odds of 1:1, on about 67% of the answers for which they gave odds of 2:1, on about 75% for odds of 3:1, etc. The actual percentages of correct answers, grouped across subjects for each of the most frequently used odds categories, are shown in the left-hand side of Table 2.

---

[1] A more detailed description of subjects' performance on this task can be found in Slovic, Fischhoff, Lichtenstein, Combs, & Layman (1976).

Table 2

Percentage of Correct Answers for Major Odds Categories

Odds:1	Experiment 1 (N=66) Lethal Events N	N%	% Correct	Experiment 3 (N=40) Lethal Events N	N%	% Correct	Experiment 4 (N=42) General Knowledge N	N%	% Correct
1	644	09	53	339	08	54	861	19	53
1.5	68	01	57	108	02.5	59	210	05	56
2	575	08	64	434	10	65	455	01	63
3	189	02	71	252	06	65	157	03.5	76
5	250	04	70	322	08	71	194	04	76
10	1,167	17	66	390	09	76	376	08	74
20	126	02	72	163	04	81	66	01.5	85
50	258	04	68	227	05	74	69	01.5	83
100	1,180	17	73	319	08	87	376	08	80
1,000	862	13	81	219	05	84	334	07	88
10,000	459	07	87	138	03	92	263	06	89
100,000	163	02	85	23	00.5	96	134	03	92
1,000,000+	157	02	90	47	01	96	360	08	94
SUM	6,098	88		2,981	68		3,855	75	
Overall Percent Correct	71.0				72.5			73.1	

Note: N% refers to the percentage of odds judgments that fell in each of the major categories. Answers pertained to frequency of lethal events in Experiments 1 and 3 and to general knowledge questions in Experiment 5. Subjects in Experiments 3 and 4 were carefully instructed about the concepts of probability, odds, and calibration.

Looking at Table 2 we see that our subjects were reasonably well calibrated at odds of 1:1, 1.5:1, 2:1, and 3:1. However, as odds increased from 3:1 to 100:1, there was little or no increase in accuracy. Only 73% of the answers assigned odds of 100:1 were correct! The accuracy percentage jumped to .81 at 1,000:1 and to .87 at 10,000:1. For the answers assigned odds of 1,000,000:1 or greater, the accuracy rate was .90. Subjects would have been well calibrated if they had assigned odds of 9:1 to the latter answers. The 12% of responses that fell between the major categories (and are not shown in Table 2) were just as poorly calibrated.

Thus, subjects in this first study exhibited extreme overconfidence. They were wrong frequently at even the highest odds levels. Moreover, they gave many extreme odds responses. Of 6,996 odds judgments, 3,560 or 51% were greater than 50:1. Almost one-fourth of the responses were greater than 1,000:1!

EXPERIMENT 2

Would the certainty illusion be exhibited by persons answering a broader range of questions and expressing their degree of certainty in probabilities rather than odds? Experiment 2 was designed to answer this question.

Method

Stimuli. The questions covered a wide variety of topics, including history, music, geography, nature, and literature. Four different formats were used. These were:

a. Open-ended format: Subjects were presented with a question stem to complete (e.g., "Absinthe is a _____."). After

writing down an answer, they estimated the probability that their answer was correct.

b. One-alternative format: Subjects were asked to estimate the probability that simple statements were correct. For example, "What is the probability that absinthe is a precious stone?" The statement of fact being judged was sometimes true, sometimes false.

c. Two-alternative format (half range of responses): For each question, subjects were asked to choose the correct answer from two which were offered. After making each choice, they judged the probability that the choice was correct. For example, "Absinthe is (A) a precious stone; (B) a liqueur." Since they chose the most likely answer, their probabilities were limited to the range [.50, 1.00].

d. Two-alternative format (full range of responses): Instead of having subjects pick the answer most likely to be correct as in Format c, the experimenters randomly selected one of the two alternatives [e.g., (B) a liqueur] and had subjects judge the probability that the selected alternative was correct. Here the full range [.00, 1.00] was used. As in Format c, one answer was correct.

Subjects and procedure. The subjects were 367 paid volunteers who responded to an ad in the University of Oregon student newspaper. They were assigned to four groups. Each group received the questions in one of the four formats. Besides the differences in question format, the specific questions used differed somewhat from group to group.

As in Experiment 1, instructions were brief and straightforward. Subjects were asked to judge the probability that an answer or statement was correct. This probability was selected from the range [.00 - 1.00] for Formats a, b, and d, and from [.50 - 1.00] for Format c.

Results[2]

Table 3 shows (a) the frequency with which subjects claimed that the probability an alternative was correct was 1.00 or .00, and (b) the percentage of answers associated with these extreme probabilities that were, in fact, correct.

The data in Table 3 tell essentially the same story as did the analysis of Experiment 1. Answers assigned a probability of 1.00 of being correct were wrong between 17% and 28% of the time. Answers assigned a probability of .00 were right between 20% and 29.5% of the time. In Formats b and d, where responses of 1.00 and .00 were possible, both responses occurred with about equal frequency. Furthermore, alternatives judged certain to be correct were wrong about as often as alternatives judged certain to be wrong were correct. The percentage of false certainties ranged from about 17% (Format a) to about 30% (Format b), but comparisons across formats should be made with caution because the item pools differed.

EXPERIMENT 3

Although the tasks and instructions for Experiments 1 and 2 seemed reasonably straightforward, we were concerned that subjects' extreme overconfidence might be due to lack of motivation or misunderstanding of the response scale. Therefore, in Experiments 3 and 4, we replicated the first two experiments, giving much more care and attention to instructing and motivating the subjects.

---

[2] Further results from this study are presented in Lichtenstein & Fischhoff (1976).

Table 3

Analysis of Certainty Responses in Experiment 2

Question Format	No. of Items	No. of Subjects	Total No. of Responses	% Certainty Responses	% Correct for Certainty Responses
a. Open-ended (p = 1.00)	43	30	1290	19.7	83.1
b. One-alternative (p = 1.00)	75	86	6450	14.2	71.7
(p = .00)				13.8	29.5
c. Two-alternative (p = 1.00)	75	120	9000	21.8	81.8
d. Two-alternative full range (p = 1.00)	50	131	6500	17.3	80.7
(p = .00)				19.1	20.5

Note: Under Formats b and d, subjects could judge the probability of the specified alternative as .00 as well as 1.00. Data from both types of certainty responses are presented here.

Method

Experiment 3 used the same 106 causes of death questions and the odds response format of Experiment 1. The experimenter started the session with a 20-minute lecture to the subjects. In this lecture, the concepts of probability and odds were carefully explained. The subtleties of expressing one's feelings of uncertainty as numerical judgments of odds were discussed, with special emphasis on how to use small odds (between 1:1 and 2:1) when one is quite uncertain about the correct answer. A chart was provided showing the relationship between various odds estimates and the corresponding probabilities. Finally, subjects were taught the concept of calibration, and were urged to make odds judgments in a way that would lead them to be well calibrated. The complete text of the instructions is presented as an appendix to this report.

The subjects for Experiment 3 were 40 persons who responded to an ad in the University of Oregon student newspaper. As in the previous experiments, they were paid for participating. Group size was held to about twenty to increase the likelihood that subjects would ask questions about any facet of the task that was unclear.

Results

The proportion of correct answers for each of the most frequent odds categories is shown in the center portion of Table 2. The detailed instructions had several effects. First, subjects were much more prone to use atypical odds such as 1.4:1, 2.5:1, etc. Only 68% of their judgments fell within the major categories of Table 2 as compared to 88% for Experiment 1. Second, their odds estimates tended to be smaller. About 43% of their estimates were 5:1 or less, compared to 27% for this category in the first

experiment. Third, subjects in this experiment were more accurate at odds above 10:1, and thus were better calibrated.

Nevertheless, subjects again exhibited unwarranted certainty. About one-third of all answers were assigned odds equal to or greater than 50:1. However, only about 83% of the answers associated with these odds were correct. When subjects estimated odds of 50:1, they were correct 74% of the time, and thus should have been giving odds of about 3:1. At 1,000:1, they should have been saying about 5:1.

Although only 68% of the responses fell in the major categories of Table 2, inclusion of the remaining 32% would not have changed the picture. Odds estimates falling between major categories were no better calibrated than estimates within those categories. We conclude that elaborate instruction of subjects tempered the certainty illusion, but only to a limited extent.

EXPERIMENT 4

Experiment 4 was similar to Experiment 3 except that subjects were asked questions dealing with topics of general knowledge (as in Experiment 2) rather than questions dealing with lethal events.

Method

The questionnaire consisted of 106 two-alternative items covering a wide variety of topics (e.g., Which magazine had the largest circulation in 1970? (A) Playboy or (B) Time; Aden was occupied in 1839 by the (A) British or (B) French; Bile pigments accumulate as a result of a condition known as (A) Gangrene or (B) Jaundice). These items were taken from a large item pool with known characteristics. Availability of this pool allowed us to select items matched in difficulty, question by question, with the 106 items about lethal events studied in Experiments 1 and 3.

The subjects were 42 paid volunteers, recruited by an ad in the University of Oregon student newspaper. The instructions paralleled those of Experiment 3. Subjects first received the detailed lecture describing the concepts of probability, odds, and calibration. They then responded to the 106 general-knowledge items, marking the answer they thought to be correct and expressing their certainty about that answer with an odds judgment. After responding to the 106 items, they were asked whether they would be willing to play a gambling game based on their odds judgments. This game is described below.

Results

The proportion of correct answers associated with each of the most common odds responses is shown on the right-hand side of Table 2. Compared with the previous studies, subjects in Experiment 4 gave a higher proportion of 1:1 odds (19% of the total responses). A few difficult items led almost all of the subjects to give answers close to 1:1, a fact indicating that they were trying to use small odds when they felt it was appropriate to do so. However, this bit of restraint was coupled with as high a percentage of large odds estimates as was given by the untutored subjects in the first experiment. About one-quarter of all answers were assigned odds equal to or greater than 1,000:1.

Once again, answers to which extremely high odds had been assigned were frequently wrong. At odds of 10:1, subjects were correct on about three out of every four questions. At 100:1, they should have been saying 4:1. At 1,000:1 and at 100,000:1, estimates of about 7:1 and 9:1 would have been more in keeping with subjects' actual abilities. Over the large number of questions for which people gave odds of 1,000,000:1 or higher, they were wrong an average of about one time out of every 16!

The gambling game. We were concerned that the odds estimates given by our subjects might not represent their real convictions. One way to test subjects' faith in their responses is to ask whether they would be willing to accept gambles contingent on the correctness of their answers and the appropriateness of their odds estimates. Given subjects' extreme overconfidence, it should be possible to construct gambles that they are eager to accept, but which, in fact, are quite disadvantageous to them.

After the subjects in Experiment 4 had answered each of the 106 questions, they were asked whether they would be willing to participate in a hypothetical game described by these instructions:

> The experiment is over. You have just earned $2.50, which you will be able to collect soon.
>
> But before you take the money and leave, I'd like you to consider whether you would be willing to play a certain game in order to possibly increase your earnings.
>
> The rules of the game are as follows:
>
> 1. Look at your answer sheet. Find the questions where you estimated the odds of your being correct as 50:1 or greater than 50:1. How many such questions were there? _____ (write number)
>
> 2. I'll give you the correct answers to these "50:1 or greater" questions. We'll count how many times your answers to these questions were wrong. Since a wrong answer in the face of such high certainty would be surprising, we'll call these wrong answers "your surprises."
>
> 3. I have a bag of poker chips in front of me. There are 100 white chips and 2 red chips in the bag. If I reach in and randomly select a chip, the odds that I will select a white

chip are 100:2 or 50:1, just like the odds that your "50:1" answers are correct.

4. For every 50:1 or greater answer you gave, I'll draw a chip out of the bag. (If you wish, you can draw the chips for me.) I'll put the chip back in the bag before I draw again, so the odds won't change. The probability of my drawing a red chip is 1/51. Since drawing a red chip is unlikely, every red chip I draw can be considered "my surprise."

5. Every time you are surprised by a wrong answer to a "50:1 or greater" question, you pay me $1. Every time I am surprised by drawing a red chip, I'll pay you $1.

6. If you are well calibrated, this game is advantageous to you. This is because I expect to lose $1 about once out of every 51 times I draw a chip, on the average. But since your odds are sometimes higher than 50:1, you expect to lose less often than that.

7. Would you play this game? Circle one. Yes No

Of the 42 subjects, 27 agreed to play. Subjects who declined were then asked if they would play if the experimenter raised the amount he would pay them to $1.50 whenever he drew a red chip, while they still had to pay only $1 in the event of a wrong answer. Six more persons agreed to play. Of the holdouts, 3 agreed to play when the experimenter offered them $2 for every red chip, and 2 more agreed when the final offer of $2.50 per red chip was made. Only three subjects refused to participate at any level of payment per red chip.

After subjects had made their decisions about playing the game, they were asked whether they would change their minds if the game were to be

played on the spot for real money. No subject indicated a desire to change his or her decision. Two subjects approached the experimenter after the experiment requesting that they be given a chance to play the game for cash. Their request was refused.

Of course, this game is strongly biased in favor of the experimenter. Since subjects were wrong about once for every eight answers assigned odds of 50:1 or greater, the game would have been approximately fair had the experimenter removed 86 of the white chips from the bag, leaving its contents at 14 white and 2 red chips.

The expected outcome of actually playing the game individually with each subject was simulated. Every wrong answer on a "50:1 or greater" question cost the subject $1. The experimenter was assumed to have drawn 1/51 of a red chip for every answer given at oods $\geq$ 50:1; his expected loss was then calculated in accordance with the bet the subject had accepted. For example, if a subject accepted the experimenter's first offer ($1 per red chip) and gave 17 "50:1 or greater" answers, the experimenter lost 17/51 dollars (33¢).

The subjects who agreed to play averaged 38.3 questions with odds $\geq$ 50:1. Thirty-six persons would have lost money and three would have won money. Individual outcomes would have ranged between a loss of $25.63 and a gain of $1.84. The mean outcome would have been a loss of $3.64 per person and the median outcome would have been a loss of $2.35. Ten persons would have lost more than $5. The 39 subjects would have lost a total of $142.13 across 1,495 answers at odds $\geq$ 50:1, an average loss of 9.5¢ for every such answer. The two persons who earnestly requested special permission to play the game would have lost $33.38 between them.

On the basis of these results, we conclude that the certainty illusion seems likely to entice people into accepting bets more disadvantageous than many that can be found in a Las Vegas casino. Indeed, it suggests that there is money to be made in "trivia-question hustling."

Our faith in this conclusion is weakened slightly by the fact that the gambles subjects accepted were hypothetical. Therefore, we ran some subjects under conditions in which they believed that they would be playing the gambling game, on the spot, for cash, with the risk of losing their own money.

We replicated Experiment 4 with 19 subjects. The only change was that the gambling game was presented as a real game. Subjects heard the instructions and decided whether or not they would play. They were told that they could lose all the money they had earned in the experiment, and possibly even more than that. After they had made their decisions, they were informed that if they earned money in the game, that amount would be added to the pay they earned for the experiment, but if they lost money, they would still leave the experiment with the money initially promised them for participating.

Six of the 19 subjects agreed to play the game as first specified (with $1 payment for each "experimenter's surprise"). Three more agreed to play when the experimenter offered to increase the payment to $1.50 per red chip. Increasing the payment to $2 brought in one additional player, and three more agreed to play at $2.50. Six subjects consistently refused to participate; some because they felt they were not well calibrated, others because they didn't like to gamble. When the game was actually played, no red chips were drawn. However, the 13 participating subjects

missed 46 of the 387 annswers to which they had assigned odds ≥ 50:1. All thirteen would have lost money, ranging from $1 to $11. Four subjects would have lost more than $6 had the game been simulated as in Experiment 4. This evidence indicates that subjects' overconfidence would lead them to fare about as poorly in a real gambling situation as in the hypothetical game described earlier.

CONCLUSION

The four experiments presented here have demonstrated the certainty illusion for questions of fact encompassing diverse content areas. Careful tutoring of subjects in the subtleties of expressing their certainty in terms of probabilities and odds did little to reduce the illusion. In attempting to explain this phenomenon, it is necessary to examine the processes of perception and memory whereby knowledge is acquired and recalled.

Nineteenth century theories of perception asserted a parallelism between mechanisms of the physical world and those of the brain. Similarly, memory was viewed as a slightly faded copy of original experience that could be retrieved accurately from some mental file. Though these views are still widely held among lay persons, psychological research over the past 80 years has led to a radically different theory of perception and memory--a theory which accommodates the certainty illusion quite comfortably.

The modern theory treats perception and memory not as copying but as decision-making processes (Neisser, 1967). According to this view, people reach conclusions about what they have seen or what they remember by reconstructing their knowledge from fragments of information, much as a paleontologist infers the appearances of a dinosaur from fragments of

bone. During reconstruction, a variety of cognitive, social and motivational factors introduce error and distortion into the output of the process.

One of the first critics of the "memory as copying" theory was Münsterberg, whose work early in this century demonstrated a form of certainty illusion in the errors of eyewitness testimony (Münsterberg, 1907; see also Loftus, 1974 and Buckhout, 1974). Münsterberg (1908) described a series of experiments demonstrating what he called "illusions of memory," unintentional mistakes made by "sound minds" in remembering details such as elapsed time, the number, speed, or size of previously observed objects, or the facts of a staged assault on a professor. He cited an experiment in which a psychologist showed three pictures to a large number of children. After observing each picture for 15 seconds, the children reported everything they could remember, underlining those details of which they were absolutely certain. According to Münsterberg, the reports contained many errors and the frequency of mistakes was almost as great among the underlined details as among the rest. Replicating these studies with adults led Münsterberg to conclude that ". . . no subjective feeling of certainty can be an objective criterion for the desired truth" (p. 490).

Research by Bartlett (1932) also contributed much to the present view of memory. Bartlett's subjects were presented with stories, prose passages, and drawings, which they were asked to reproduce after varying intervals of time. Bartlett observed that accurate recall of the material was rare--that reconstruction and change were the rule rather than the exception. He found, for example, that subjects not only created new

material but were often most certain about that which they had invented.

Treating perception and memory as inferences focuses attention on the logical processes by which fragments of information are molded into conclusions about the world. People appear to be insufficiently critical of the information, assumptions, and reasoning they use to reconstruct their knowledge, and this in turn leads to overconfidence. Consider, for example, the work of Johnson-Abercrombie (1960) who studied the responses of medical students shown x-rays of two hands and asked "to list the differences between the two hands." The students' observations fell into two categories. The first consisted of simple descriptive statements (facts) about differences in size, number, shape, and distribution of the shadows in the prints. The second, and larger, category included inferences such as "A is a young hand and B is an old hand." Often the students appeared not to recognize the difference between facts and inferences, nor the assumptions on which their inferences were based. For example, the inference about age appeared to be based on the smaller size of Hand A and greater number of bones in it. The students had taken for granted that the size of the prints was a sure guide to the size of the hands themselves (an assumption that was actually quite tenuous). Similarly, although the number of bones might be a cue to age, many other reasons for the differences were overlooked--for example, the possibility that in B the bones overlapped or had been resorbed or that the "hands" came from different species of animals. Johnson-Abercrombie concluded that:

> The inferences the students had made were not arrived at as a series of logical steps but swiftly and almost unconsciously. The validity of the inferences was usually

not inquired into; indeed, the process was usually accompanied by a feeling of certainty of being right . . . (p. 89).[3]

Examination of the errors made by subjects in the present experiments provides further insight into the nature and pitfalls of the processes whereby knowledge is reconstructed. Consider first the errors in judgments about the relative frequencies of lethal events in Experiment 3. Tversky and Kahneman (1973) have proposed that people's judgments of event frequency are inferences, constructed according to the ease with which relevant instances of the event can be imagined or by the number of such instances that are readily retrieved from memory. They called the use of imaginability and memorability as cues for frequency the "availability" mechanism. According to this mechanism, one's direct experiences with a

---

[3] Another example of the subtle role of assumptions in the reconstruction of knowledge, this time applied to memory, comes from the experience of one of the authors who became embroiled in a friendly debate with a colleague about the dates of a forthcoming conference. Both parties agreed that the conference was to last about 4-5 days. But the dispute centered about whether these dates were March 30-April 3 or April 30-May 3. The author was certain of the former dates because he specifically recalled the date March 30 in the organizer's letter. His colleague was certain of the latter period because he specifically recalled the date May 3 in the letter. Bets were placed and the letter was consulted to resolve the dispute. To the surprise of both parties, the letter stated the dates as March 30-May 3, an obvious mistake. Thus, both paries were correct regarding the fragment of information they recalled, but one fragment led to the wrong conclusion.

lethal event should affect one's judgments of its frequency. One's indirect exposure to the event via movies, books, newspaper publicity, etc. should also influence judged frequency. In Experiments 1 and 3, many of the items about which subjects were certain but wrong can be attributed to the overestimation of dramatic, well-publicized (i.e., readily available) events such as death from homicide, accidents, pregnancy and abortion, and the underestimation of "quiet" killers such as emphysema, diabetes, and appendicitis. For example, about 30% of the subjects gave odds $\geq$ 50:1 that homicide was more frequent than suicide. Actually, suicide takes about 25% more lives each year.

Subjects in Experiment 3 were asked to select one answer about which they were certain and to write a short statement indicating why they were so confident. One subject described his odds of 2,000:1 that deaths from pregnancy, birth and abortion were more frequent than deaths from appendicitis by writing, "I've never heard of a person dying of appendicitis, but I have many times heard of persons dying during childbirth and abortion." The availability mechanism is generally a useful device for inferring frequency. After all, our everyday experience has taught us that frequent events often are easier to imagine and recall than infrequent events. However, availability is also affected by recency, emotional saliency, and other factors that may be unrelated to actual frequency. Unless one's degree of certainty can be tuned to these factors, it may be biased.

An interesting error, probably indicative of something other than availability bias, was subjects' tendency to judge incorrectly, yet with certainty, that death from smallpox was more frequent than death from smallpox vaccination. Perhaps the subjects were relying on the generally valid assumption that vaccines are safer than the diseases they are meant

to prevent. In this case, that assumption is misleading. The vaccine has been so successful that there has not been a death in the U. S. from smallpox since 1949, while 6-10 persons die annually from complications arising from the vaccination.

Table 4 presents several items from Experiment 4 which are suggestive of the sorts of inferences that led people astray when judging general knowledge questions. When answering Question 1, many subjects may have assumed that the reference was to the first airplane raid, whereas it really referred to any sort of airborn attack, the first on record being Austria's use of balloons to bomb Venice in 1849. Regarding Question 2, the importance of the potato in Irish history may have led subjects to believe that Ireland was equally important for the potato's development. Regarding Question 3, cacao, like the potato, is native to South America, a fact that subjects may have known (or guessed from its Spanish-sounding name). The subjects may have been misled by assuming that cacao's continent of origin is also the continent where production is greatest. Subjects were probably wrong about Question 4 because they drew the inference that Adonis was the God of Love from the fact that he was a handsome youth who had an affair with Venus, the Goddess of Love. The error made about Time versus Playboy (Question 5) possibly arose from the fact that Time is a worldwide magazine that has led its field for more than 30 years, whereas Playboy is a much newer and more specialized publication; it is also possible that readers of Time are more likely to leave the magazine lying about visibly and make reference to it in conversation. Although these comments about specific items are speculative, they may illuminate the workings of the inferential processes that underlie our knowledge about our knowledge.

Table 4

Odds Responses to Five High Certainty-Low Accuracy Items in Experiment 4

1. The first air raid in history took place in (A) 1849; (B) 1937?
2. Potatoes are native to (A) Ireland; (B) Peru?
3. Three-fourths of the world's cacao comes from (A) Africa; (B) South America?
4. Adonis was the Semitic god of (A) love; (B) vegetation?
5. Which magazine had the larger circulation in 1970? (A) Playboy; (B) Time

1		2		3		4		5	
(A)	1849	A.	Ireland	(A)	Africa	A.	love	(A)	Playboy
B.	1937	(B)	Peru	B.	South America	(B)	vegetation	B.	Time
A.	50	B.	100,000	A.	5,000	B.	1,000	A.	10,000
A.	10	B.	100,000	A.	1	B.	15	A.	100
A.	3	B.	10,000	B.	1	B.	9.5	A.	50
A.	3	B.	1,000	B.	1.2	B.	5	A.	10
A.	2	B.	1,000	B.	1.5	B.	3	A.	10
A.	2	B.	100	B.	1.9	B.	3	A.	9
A.	1.5	B.	100	B.	2	B.	2	A.	7
A.	1.2	B.	100	B.	2	B.	1.1	A.	2
A.	1.1	B.	100	B.	2	B.	1	A.	2
A.	1	B.	40	B.	2	B.	1	A.	2
A.	1	B.	6	B.	2	B.	1	A.	2
B.	1	B.	5	B.	2	B.	1	A.	2
B.	1	B.	3	B.	3	B.	1	A.	2
B.	1	B.	2	B.	3	A.	1	A.	1.5
B.	1.5	B.	1.5	B.	3	A.	1	A.	1.5
B.	1.5	A.	1	B.	3	A.	1	B.	1
B.	2	A.	1	B.	3	A.	1.1	B.	1
B.	2	A.	1	B.	5	A.	1.1	B.	1
B.	2	A.	1.5	B.	5	A.	1.5	B.	1
B.	2	A.	2	B.	10	A.	1.5	B.	1.1
B.	2	A.	2	B.	10	A.	1.5	B.	1.5
B.	3	A.	2	B.	10	A.	1.5	B.	2
B.	3	A.	2	B.	10	A.	1.6	B.	2
B.	5	A.	3	B.	10	A.	2	B.	2
B.	10	A.	5	B.	10	A.	2	B.	2
B.	10	A.	5	B.	25	A.	2	B.	2
B.	10	A.	5	B.	30	A.	2	B.	5
B.	10	A.	10	B.	50	A.	3	B.	5
B.	20	A.	10	B.	100	A.	4	B.	5
B.	20	A.	10	B.	100	A.	10	B.	5
B.	20	A.	20	B.	100	A.	20	B.	10
B.	25	A.	25	B.	100	A.	30	B.	10
B.	100	A.	75	B.	100	A.	100	B.	10
B.	100	A.	100	B.	500	A.	100	B.	10
B.	100	A.	100	B.	500	A.	100	B.	40
B.	1,000	A.	100	B.	1,000	A.	100	B.	100
B.	1,000	A.	100	B.	1,000	A.	100	B.	100
B.	5,000	A.	1,000	B.	1,000	A.	1,000	B.	500
B.	10,000	A.	1,000	B.	10,000	A.	4,000	B.	1,000
B.	1,000,000	A.	1,000	B.	10,000	A.	10,000	B.	1,000
B.	1,000,000	A.	1,000,000	B.	1,000,000	A.	10,000	B.	10,000
B.	1,000,000	A.	1,000,000	B.	1,000,000	A.	1,000,000	B.	100,000

Note: Correct answers are circled. Answers below the horizontal line are incorrect.

Besides being an interesting psychological phenomenon, the certainty illusion may have important social consequences. Münsterberg (1908) bemoaned the fact that, every day, in thousands of courts across the world, witnesses under oath affirmed mixtures of truth and untruth, memory and illusion, knowledge and suggestion, experience and wrong conclusions. He stressed the contribution that psychology could make to improving the legal system.

> It is in normal mental life . . . that the progress of psychological science cannot be further ignored. No railroad or ship company would appoint to a responsible post . . . men whose eyesight had not been tested for colour blindness.
>
> In the life of justice, trains are wrecked and ships are colliding too often, simply because the law does not care to examine the mental colour blindness of the witness's memory (pp. 68-69).

Almost 70 years later, the trains and ships have moved aside to make room for planes and missiles, nuclear weapons and reactors, and other creations of advanced technology. More and more, we are asked to place our trust in the certainty of expert judgment (Fischhoff, 1976). Münsterberg's admonition seems no less relevant today.

APPENDIX: DETAILED INSTRUCTIONS FOR EXPERIMENTS 3 AND 4

You're probably all somewhat familiar with the notions of probability odds, and chance, and that's what this experiment is all about. We're interested in your ability to translate your own feelings of uncertainty into numerical judgments of odds.

During the main part of the experiment, you will be given a multiple choice test. For each question, you will be asked first to select which of two alternatives, A or B, is correct and then to express your sureness in your answer in terms of the odds, that your answer is correct.

We'll be looking to see how accurate your odds judgments are. If they are accurate, you can consider yourself a "well calibrated odds assessor" and add this skill to your other accomplishments.

Because we have found that these judgments are not always easy to make, I'd first like to spend some time discussing the concepts of probability and odds with you and explaining what you have to do to be "well calibrated."

First, let me clarify my use of the words "probability" and "odds." A probability is a number between 0 and 1, such as .24 or .5 or .8333, which expresses a degree of certainty or uncertainty. Odds are two numbers, written with a colon in between, such as 1:2 or 6:1. This pair of numbers also expresses a degree of uncertainty, and there is a regular relationship between any probability number and an odds. Saying "the probability is .5" means exactly the same as saying "the odds are 1:1." Whenever the probability is greater than .5, like .75, then the odds' first number is larger than the second number. "The probability is .75" means exactly the same as "the odds are 3:1." When the probability is less than .5, the odds' first number is smaller than the second.

Example: "The probability is .25" = "The odds are 1:3."

In this experiment, we will be asking you for odds, not probabilities. And these odds will always be of the form "more likely than not"--odds where the chosen alternative is more likely true than not true. So the odds' first number will always be larger than the odds' second number. In particular, we will ask you for odds in which the second number is always one. All the odds will be of the form:

X : 1

For those of you who are more familiar with probabilities than with odds, we have prepared a chart showing you the relationship between odds and probabilities. You can keep this chart handy if you wish to consult it.

Now we'll talk about how to do it. Suppose I give you this test item:

Tomorrow it will

A. rain
B. not rain

Today, you are not sure of the answer. You are uncertain. The question is, how uncertain? If you think it is more likely to rain than not to rain, you should pick alternative A. Then you will have to choose a number to express your degree of certainty, in odds. Suppose your odds response is

10 : 1

That means that you think the chances of rain tomorrow are 10 times as great as the chances of no rain. If you say

100 : 1

you think the chances of rain tomorrow are 100 times more likely than the chances of no rain. With the odds of 100:1, you are much more certain that your chosen alternative (rain tomorrow) is correct than with the odds of 10:1.

But suppose that you are completely uncertain--maybe it'll rain and maybe it won't, but you wouldn't be willing to bet more on one alternative than the other. There is a special odds for this situation:

1 : 1

When you give odds of 1:1, you are saying that both alternatives are equally likely, you couldn't possibly pick between them.

Now suppose that you think rain is just a little bit more likely than no rain. Not twice as likely (odds = 2:1), but less than twice as likely. Then you must use a decimal point in your odds, giving some answer like:

1.5 : 1

Odds of 1.5:1 mean that you think the chosen alternative is just one and a half times as likely as no rain.

One way of understanding these small odds is by translating them into probabilities:

Odds of 1.0:1 =	probability of .50
1.1:1 =	.52
1.2:1 =	.55
1.3:1 =	.57
1.4:1 =	.58
1.5:1 =	.60
1.6:1 =	.62
1.7:1 =	.63
1.8:1 =	.64
1.9:1 =	.66
2.0:1 =	.67
2.5:1 =	.71
3.0:1 =	.75
3.5:1 =	.78
4.0:1 =	.80
etc.	

(These are not exact translations; we have rounded the probabilities to the nearest hundredth.)

In other words, saying "the odds are 1.3:1" is the same as saying "the probability that it will rain tomorrow is .57 (while the probability it will not rain is .43)."

The items we are using today are not items about future events, but items of fact, and in every case, we know the correct answer.

You may feel certain, or almost certain, that you know the correct answer, too. If you do, it is appropriate for you to report very large odds, such as

100 : 1
125 : 1
500 : 1
1000 : 1

. . . or even larger.

Saying 10,000:1 means that you feel it is ten thousand times more likely that you are right than that you are wrong. Of course, you may be quite uncertain about the answer. In that case, your opinion may be much better represented by small odds, perhaps as low as 1:1.

Summary so far:

1. You will express your confidence that you have chosen the correct alternative in the form of odds,

X : 1

When X is some number equal to or larger than one.

2. "1:1" means you are just as likely to be wrong as to be right.

3. The more certain you are to be right, the larger the number you should choose.

4. You may choose any number you wish (as long as the number is equal to or larger than one), including decimal numbers such as 1.3:1. Odds less than 2:1 are especially useful when you are quite uncertain, that is, when your chances of being right are less than twice as large as your chances of being wrong.

But what number should you choose? This is the n[illegible] the problem. We are asking you to do a very difficult task. We want y[illegible]o examine your own "gut feelings" of certainty and uncertainty, and translate those feelings into an odds number. There are no rules for doing this.

There are, however, two criteria to keep in mind. Knowing these may help you in your task.

First, your odds should reflect how much you actually know. That is, your odds should be as extreme as your knowledge justifies. If you are indeed correct, odds of 10,000:1 are better than odds of 10:1.

But extremeness is not the only requirement. The second is called calibration.

Consider all the times in this experiment that you say "the odds are 10:1." Now in addition, consider all the times in your whole life that you say "the odds are 10:1 that I'm right." For all these many occasions, suppose we can later determine whether you were right or wrong. This one characteristic of all those occasions should be true:

For every 10 times you were right, you should be wrong once.

In other words, when you say

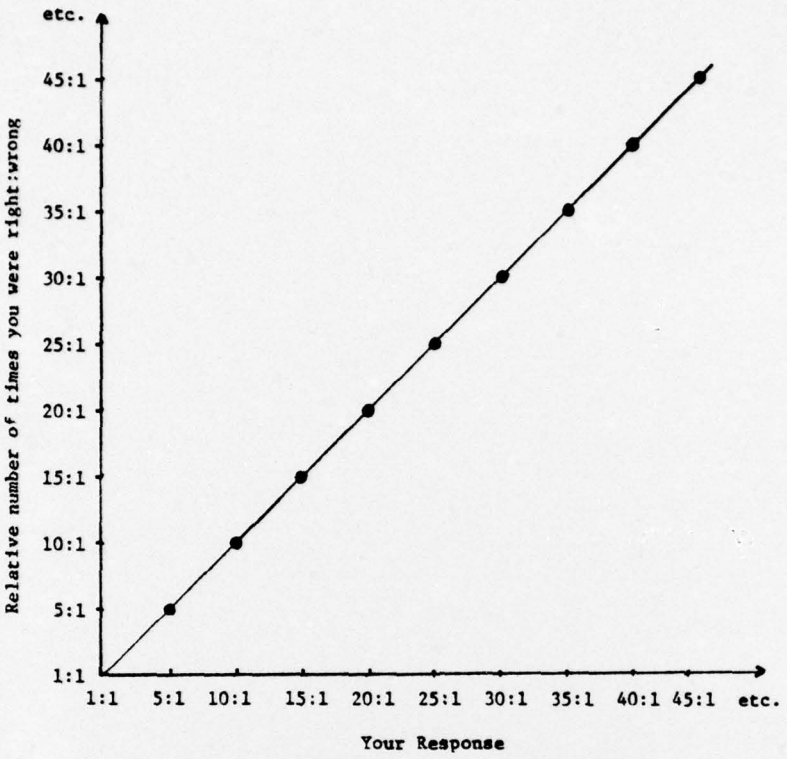
10 : 1

you are saying--"I'm not absolutely sure I'm right, I think I'll be right just 10 times out of every 11, and wrong 1 time out of every 11." The number you give should match the frequency, over many occasions on which you said the same thing, that you're right. "The odds are 2:1" means: "Over the hundreds, thousands, of times I say, or think, that the odds are 2:1, on the average, I should be right two times out of three, and wrong one third of the time."

When this match between what you say and how often you're right occurs, we say that you are "well calibrated." People are not always well calibrated. Sometimes they are right more often than the odds they report would lead us to expect, and sometimes they are right less often than the odds suggest.

Your task in this experiment is to choose the odds that are as precisely calibrated as possible.

We will take your answers and check your calibration. We will group together all items for which you said 2:1. In another group will be all the items for which you said 10:1. And so forth. In each group, we will observe the relative number of times you were right or wrong. We will plot this as follows:



Of course, even if you were perfectly calibrated, the chart might not look so perfect, because--maybe you'll say a certain odds just 15 times, and even if, on the average, your frequency of being correct is right on, a sample of 15 might be a bit off.

But don't worry about that. Analyzing the data is our problem. Your problem is to try, every time you make an odds response (say X:1), to give a response such that, in the long run, for every X plus one occasions, you'll be right X times and wrong once.

I realize that I <u>still</u> haven't told you how to arrive at the number you use in your odds. I've tried. But when it comes right down to it, I can't. I can (and I hope I just did) explain the <u>meaning</u> of odds. But you are the only one who can tell, for a given item, just how uncertain you are about the correct answer to that item. You are the only judge of your feelings and your beliefs. It is your job to translate those gut feelings and beliefs into odds.

A few final cautions: In these instructions and explanations we have relied on just a few odds to use as examples: 2:1, 10:1, etc., for simplicity. Do not suppose that you have to use those odds, or that you should avoid using those odds. Use any odds you want to use, so long as the odds you choose are equal to or greater than one. Use whatever odds best express your feelings of uncertainty.

The chart we have given you is designed to help those people to "think better" in units of probabilities than in units of odds. Its only function

is to help you out if you find it helpful. You can disregard it if you wish. Do not feel limited by the numbers in the chart. If you really believe that some other number--like 1.05:1 or 15:1, or whatever--is appropriate, then use it.

On the other hand, the task is hard enough as it is. Don't torture yourself trying to decide whether to answer 15:1 or 16:1. Nobody's good enough at knowing their own uncertainty to make a subtle discrimination like that. Try hard to be thoughtful and to be careful, but don't tie yourself into knots.

Complete every item; try not to miss any. If you have a change of heart, you can, and should, go back and change an answer.

As I said before, if you don't have any idea at all about which alternative is correct, so that you would be willing to let the flip of a coin choose the alternative, you should give the response:

1 : 1

Are there any questions?

[Specific instructions for the lethal events or general knowledge tasks then followed.]

REFERENCES

Bartlett, F. C. Remembering. Cambridge, England: Cambridge University Press, 1932.

Buckhout, R. Eyewitness testimony. Scientific American, 1974, 231, 23-31.

Fischhoff, B. Cost-benefit analysis and the art of motorcycle maintenance. Oregon Research Institute Research Monograph, 1976, 16, 1.

Johnson-Abercrombie, M. L. The anatomy of judgment. New York: Basic Books, 1960.

Lichtenstein, S., & Fischhoff, B. Do those who know more know more about how much they know? The calibration of probability judgment. Oregon Research Institute Research Bulletin, 1976, 16, 1.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. Calibration of probabilities: The state of the art. In H. Jungermann and G. de Zeeuw (Eds.), Proceedings of the Fifth Conference on Subjective Probability, Utility, and Decision Making. Darmstadt, Germany, in press.

Loftus, E. The incredible eyewitness. Psychology Today, 1974, 8, 116-119.

Münsterberg, H. On the witness stand. New York: Doubleday, Page & Co., 1908.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

Slovic, P., Fischhoff, B., Lichtenstein, S., Combs, B., & Layman, M. Perceived frequency of low probability, lethal events. Oregon Research Institute Research Bulletin, 1976, 16, 5.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 1973, 5, 207-232.

Research Distribution List

Department of Defense

Assistant Director (Environment and Life Sciences)
Office of the Deputy Director of Defense Research and Engineering (Research and Advanced Technology)
Attention: Lt. Col. Henry L. Taylor
The Pentagon, Room 3D129
Washington, DC 20301

Office of the Assistant Secretary of Defense (Intelligence)
Attention: CDR Richard Schlaff
The Pentagon, Room 3E279
Washington, DC 20301

Director, Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Director, Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Director, Program Management Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209
(two copies)

Administrator, Defense Documentation Center
Attention: DDC-TC
Cameron Station
Alexandria, VA 22314
(12 copies)

Department of the Navy

Office of the Chief of Naval Operations (OP-987)
Attention: Dr. Robert G. Smith
Washington, DC 20350

Director, Engineering Psychology Programs (Code 455)
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217
(three copies)

Assistant Chief for Technology (Code 200)
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Office of Naval Research (Code 230)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Naval Analysis Programs (Code 431)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Operations Research Programs (Code 434)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research (Code 436)
Attention: Dr. Bruce McDonald
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Information Systems Program (Code 437)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research (ONR)
International Programs (Code 1021P)
800 North Quincy Street
Arlington, VA 22217

Director, ONR Branch Office
Attention: Dr. Charles Davis
536 South Clark Street
Chicago, IL 60605

Director, ONR Branch Office
Attention: Dr. J. Lester
495 Summer Street
Boston, MA 02210

Director, ONR Branch Office
Attention: Dr. E. Gloye and Mr. R. Lawson
1030 East Green Street
Pasadena, CA 91106
(two copies)

Dr. M. Bertin
Office of Naval Research
Scientific Liaison Group
American Embassy – Room A-407
APO San Francisco 96503

Director, Naval Research Laboratory
Technical Information Division (Code 2627)
Washington, DC 20375
(six copies)

Director, Naval Research Laboratory (Code 2029)
Washington, DC 20375
(six copies)

Scientific Advisor
Office of the Deputy Chief of Staff
for Research, Development and Studies
Headquarters, U.S. Marine Corps
Arlington Annex, Columbia Pike
Arlington, VA 20380

Headquarters, Naval Material Command (Code 0331)
Attention: Dr. Heber G. Moore
Washington, DC 20360

Headquarters, Naval Material Command (Code 0344)
Attention: Mr. Arnold Rubinstein
Washington, DC 20360

Naval Medical Research and Development Command (Code 44)
Naval Medical Center
Attention: CDR Paul Nelson
Bethesda, MD 20014

Head, Human Factors Division
Naval Electronics Laboratory Center
Attention: Mr. Richard Coburn
San Diego, CA 92152

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Naval Personnel Research and Development Center
Management Support Department (Code 210)
San Diego, CA 92152

Naval Personnel Research and Development Center (Code 305)
Attention: Dr. Charles Gettys
San Diego, CA 92152

Dr. Fred Muckler
Manned Systems Design, Code 311
Navy Personnel Research and Development Center
San Diego, CA 92152

Human Factors Department (Code N215)
Naval Training Equipment Center
Orlando, FL 32813

Training Analysis and Evaluation Group
Naval Training Equipment Center
(Code N-00T)
Attention: Dr. Alfred F. Smode
Orlando, FL 32813

Department of the Army

Technical Director, U.S. Army Institute for the Behavioral and Social Sciences
Attention: Dr. J.E. Uhlaner
1300 Wilson Boulevard
Arlington, VA 22209

Director, Individual Training and Performance Research Laboratory
U.S. Army Institute for the Behavioral and
and Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

Director, Organization and Systems Research Laboratory
U.S. Army Institute for the Behavioral and
Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

Department of the Air Force

Air Force Office of Scientific Research
Life Sciences Directorate
Building 410, Bolling AFB
Washington, DC 20332

Robert G. Gough, Major, USAF
Associate Professor
Department of Economics, Geography and
Management
USAF Academy, CO 80840

Chief, Systems Effectiveness Branch
Human Engineering Division
Attention: Dr. Donald A. Topmiller
Wright-Patterson AFB, OH 45433

Aerospace Medical Division (Code RDH)
Attention: Lt. Col. John Courtright
Brooks AFB, TX 78235

Other Institutions

The Johns Hopkins University
Department of Psychology
Attention: Dr. Alphonse Chapanis
Charles and 34th Streets
Baltimore, MD 21218

Institute for Defense Analyses
Attention: Dr. Jesse Orlansky
400 Army Navy Drive
Arlington, VA 22202

Director, Social Science Research Institute
University of Southern California
Attention: Dr. Ward Edwards
Los Angeles, CA 90007

Perceptronics, Incorporated
Attention: Dr. Amos Freedy
6271 Variel Avenue
Woodland Hills, CA 91364

Director, Human Factors Wing
Defense and Civil Institute of
Environmental Medicine
P.O. Box 2000
Downsville, Toronto
Ontario, Canada

Stanford University
Attention: Dr. R.A. Howard
Stanford, CA 94305

Montgomery College
Department of Psychology
Attention: Dr. Victor Fields
Rockville, MD 20850

General Research Corporation
Attention: Mr. George Pugh
7655 Old Springhouse Road
McLean, VA 22101

Oceanautics, Incorporated
Attention: Dr. W.S. Vaughan
3308 Dodge Park Road
Landover, MD 20785

Director, Applied Psychology Unit
Medical Research Council
Attention: Dr. A.D. Baddeley
15 Chaucer Road
Cambridge, CB 2EF
England

Department of Psychology
Catholic University
Attention: Dr. Bruce M. Ross
Washington, DC 20017

Stanford Research Institute
Decision Analysis Group
Attention: Dr. Allan C. Miller III
Menlo Park, CA 94025

Human Factors Research, Incorporated
Santa Barbara Research Park
Attention: Dr. Robert R. Mackie
6780 Cortona Drive
Goleta, CA 93017

University of Washington
Department of Psychology
Attention: Dr. Lee Roy Beach
Seattle, WA 98195

Eclectech Associates, Incorporated
Post Office Box 179
Attention: Mr. Alan J. Pesch
North Stonington, CT 06359

Hebrew University
Department of Psychology
Attention: Dr. Amos Tversky
Jerusalem, Israel

Dr. T. Owen Jacobs
Post Office Box 3122
Ft. Leavenworth, KS 66027

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER / 2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) The Certainty Illusion. (9)	5. TYPE OF REPORT & PERIOD COVERED Technical Rept.
	6. PERFORMING ORG. REPORT NUMBER ORI Report No.: DDI-2
7. AUTHOR(s) (10) Paul Slovic, Baruch Fischhoff Sarah Lichtenstein	8. CONTRACT OR GRANT NUMBER(s) Prime Contract No.: N00014-76-C-0074 Subcontract No.: 75-030-0712
9. PERFORMING ORGANIZATION NAME AND ADDRESS Oregon Research Institute P. O. Box 3196 Eugene, Oregon 97401	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, Virginia 22209 (11)	12. REPORT DATE August 1976 13. NUMBER OF PAGES 41
14. MONITORING AGENCY NAME & ADDRESS (If different from Controlling Office) Office of Naval Research 800 North Quincy Street Arlington, VA 22217 (12) 38 p.	15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

(15) N00014-76-C-0074,

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

√√ ARPA Order-3052

18. SUPPLEMENTARY NOTES

Support for this research performed by Oregon Research Institute was provided by the Advanced Research Projects Agency of the Department of Defense and was monitored under Contract N00014-76-C-0074 with the Office of Naval Research under subcontract from Decisions and Designs, Inc.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Perceived probability
Decision making
Overconfidence
Judgment

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

When we feel certain about our factual knowledge, all too often we are wrong. This phenomenon, labeled "the certainty illusion," is demonstated in four experiments in which subjects (1) answered questions about a variety of topics and (2) indicated their degree of certainty about each answer. Subjects were wrong frequently on answers judged certain to be correct. Careful tutoring of subjects in the subtleties of expressing their certainty in terms of probabilities and odds did little to reduce the illusion. Feelings of certainty were so strong that subjects were willing to bet on the correctness of their knowledge. Because of the illusion,

continued (on p. 36)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Abstract continued:

the bets they accepted were quite disadvantageous to them. The psychological basis for unwarranted certainty is discussed in terms of the inferential processes whereby knowledge is reconstructed from fragments of perceptions and memories.